FACULTY OF **ARTS AND PHILOSOPHY**

Koen Plevoets, Bart Defrancq

INFORMATIONAL LOAD AS A TRIGGER FOR DISFLUENCIES IN INTERPRETING

Overview

Intro

Research questions

Data

Method

Analysis

Conclusion


EUROPEAN PARLIAMENT

Intro

Interpreting: The rendition of utterances in another language

- Consecutive
- Simultaneous
- …

'Real-time' translation

Intro

‘Interpreting is a cognitively demanding activity’

- Multitasking: Division of attention to different concurring tasks
- ‘Tightrope hypothesis’: Interpreters work at the limits of their processing capacities

(Gile 1999)

Intro

Effort Model

(Gile 1985; 1997)

Interpreting= L + P + M + C

- L: Listening effort
- P: Production effort
- M: Memory effort
- C: Coordination effort

Intro

Cognitive Load Model

(Seeber 2011; 2013)

Interpreting: “Real-time combination of language comprehension and language production task”

Both tasks have “demand vectors”:

- Auditory verbal
- Cognitive-verbal
- Verbal-response (only for production)

Gerver (1976)

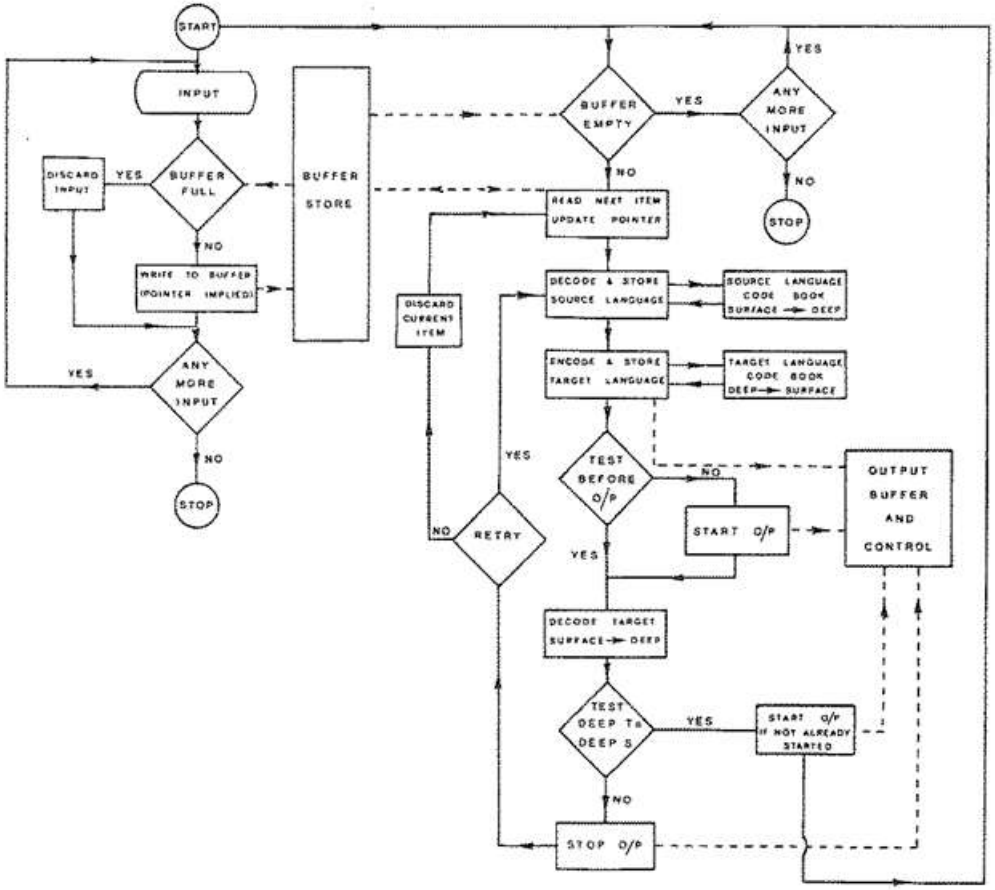
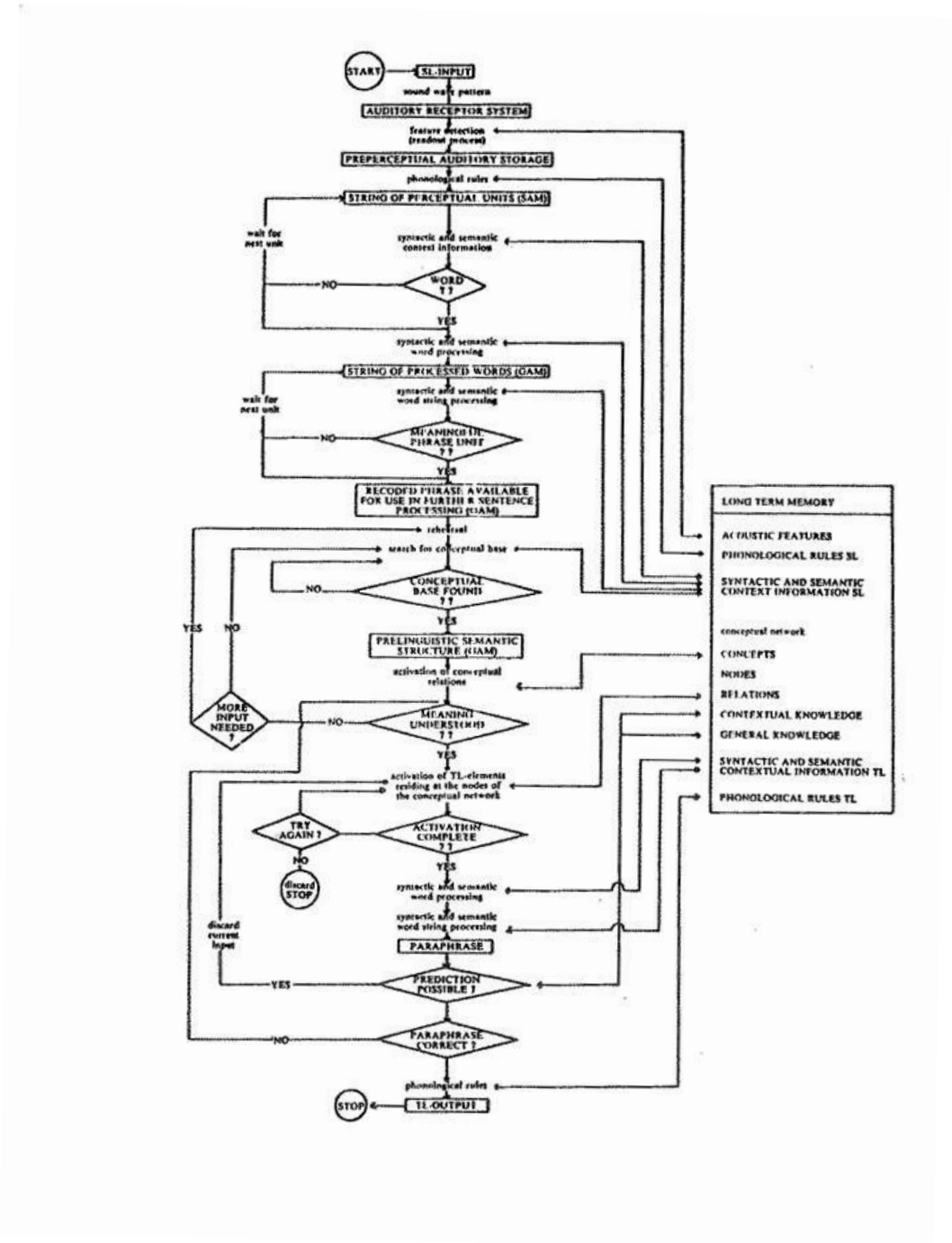
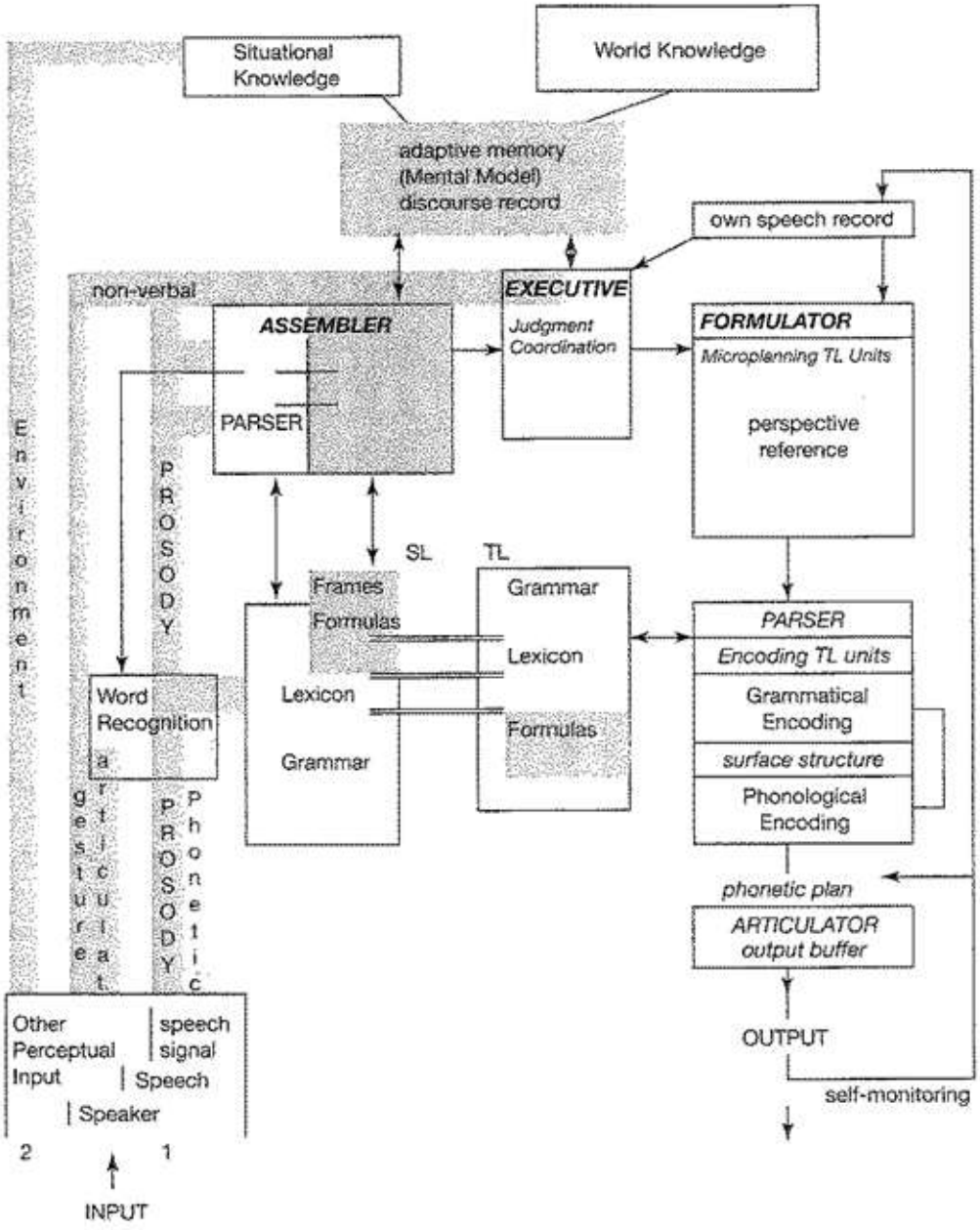


Figure 1: A model of the simultaneous interpretation process.

Moser (1978)


START
SL-INPUT
sound wave pattern
AUDITORY RECEPTOR SYSTEM
feature detection (readout process)
PREPERCEPTUAL AUDITORY STORAGE
phonological rules
STRING OF PERCEPTUAL UNITS (SAM)
wait for next unit
syntactic and semantic context information
NO
WORD ??
YES
syntactic and semantic word processing
STRING OF PROCESSED WORDS (SAM)
wait for next unit
syntactic and semantic word string processing
NO
MEANINGFUL PHRASE UNIT ??
YES
RECODED PHRASE AVAILABLE FOR USE IN FURTHER SENTENCE PROCESSING (SAM)
rehearsal
search for conceptual base
NO
CONCEPTUAL BASE FOUND ??
YES
YES
NO
PRELINGUISTIC SEMANTIC STRUCTURE (SAM)
activation of conceptual relations
MORE INPUT NEEDED ?
NO
MEANING UNDERSTOOD ??
YES
activation of TL-elements residing at the nodes of the conceptual network
TRY AGAIN ?
ACTIVATION COMPLETE ??
NO
YES
discard STOP
syntactic and semantic word processing
discard current input
syntactic and semantic word string processing
PARAPHRASE
YES
PREDICTION POSSIBLE ?
NO
PARAPHRASE CORRECT ?
phonological rules
STOP
TL-OUTPUT
LONG TERM MEMORY
ACOUSTIC FEATURES
PHONOLOGICAL RULES SL
SYNTACTIC AND SEMANTIC CONTEXT INFORMATION SL
conceptual network
CONCEPTS
NODES
RELATIONS
CONTEXTUAL KNOWLEDGE
GENERAL KNOWLEDGE
SYNTACTIC AND SEMANTIC CONTEXTUAL INFORMATION TL
PHONOLOGICAL RULES TL

Setton (1999)


Situational Knowledge
World Knowledge
adaptive memory (Mental Model) discourse record
own speech record
non-verbal
EXECUTIVE
Judgment Coordination
ASSEMBLER
PARSER
FORMULATOR
Microplanning TL Units
perspective reference
Environment
PROSODY
SL
TL
Frames
Formulas
Lexicon
Grammar
Grammar
Lexicon
Formulas
PARSER
Encoding TL units
Grammatical Encoding
surface structure
Phonological Encoding
Word Recognition
gestural
articulat
PROSODY
Phonetic
phonetic plan
ARTICULATOR
output buffer
OUTPUT
self-monitoring
Other Perceptual Input
speech signal
Speech
Speaker
2
1
INPUT

Intro

Information overload unsettles interpreting

Errors and omissions

- Delivery rate (Gerver 1969; Pio 2003)
- Propositional density (Dillinger 1994; Tommola & Helevä 1998)

BUT: Vague definition of ‘interpreting error’

(Barik 1975; Gerver 1976)

Intro

Information overload unsettles interpreting

Disfluencies: Pauses, *uh(m)*,…

(Goldman-Eisler 1967; Mead 2000; Tissi 2000; Cecot 2001)

Vast psycholinguistic literature

(Levelt 1983; Arnold et al. 2000; 2003; Bortfeld et al. 2001; Clark & Fox Tree 2002; Watanabe et al. 2008,…)

Intro

Setton (1999: 247)

	Attention to input	**Attention to formulation**
Long silent pause	High	-
Short pausing	Normal listening	Routine planning
Filled pause	Normal listening	Routine planning
Mixed: Short & filled pauses & voice effects	Normal listening	Routine planning
Long filled pause	Relaxed or off	Planning/Searching
Fluent unmodulated string	Relaxed or off	Off

Intro

Naturalistic data: Corpus-based

(Gile 1998)

Next to classical study of source influence on target, also ‘Bakerian’ comparison of interpreting with non-interpreting

(Baker 1993)

Research questions

1. How does informational load manifest itself in interpreting vs. non-interpreting?
2. How does the input load differ from the output load in interpreting?

Data

2 corpora:

- European Parliament Interpreting Corpus – Ghent
- Spoken Dutch Corpus – component g

Data

European Parliament Interpreting Corpus – Ghent

Plenary sessions of the European Parliament

2006-2008

French, Spanish, Dutch, and English

190 000 tokens… and rising

Data

European Parliament Interpreting Corpus – Ghent

Transcribed according to VALIBEL-corpus (Bachy et al. 2007)

POS-tagged and chunked by means of LeTs (Van de Kauter et al. 2013)

Sentence-aligned with WinAlign

Data

Spoken Dutch Corpus – component g
(Oostdijk 2000)

Parliamentary debates

1998-2003

POS-tagged

Data

Spoken Dutch Corpus – component g

(Oostdijk 2000)

360 000 tokens

- Flanders: 140 000
- The Netherlands: 220 000

Data

		Nr. of files	Nr. of sentences
EPICg	FRA (source)	108	1458
	DUT (target)	108	1437
SPCg		240	19046
	(Fl	155	8293)
	(Nl	85	10753)

Method

Predict the nr. of *uh(m)*'s in each sentence on the basis of:

- Lexical density
- Proportion of numbers (Gile 2009)
- Delivery rate

Method

Lexical density: Nr. of content words / (nr. of content words + nr. of function words)

Proportion of numbers: Nr. of numerals / total nr. of words

Delivery rate: Total nr. of words / total nr. of minutes

Method

Multilevel rate model

(Faraway 2006: 61-63, 221-230)

Predict nr. of *uh(m)*'s per sentence, conditioned on the sentence's total nr. of words ('offset')

Random factor: files (108+240)

Method

First-level predictors:

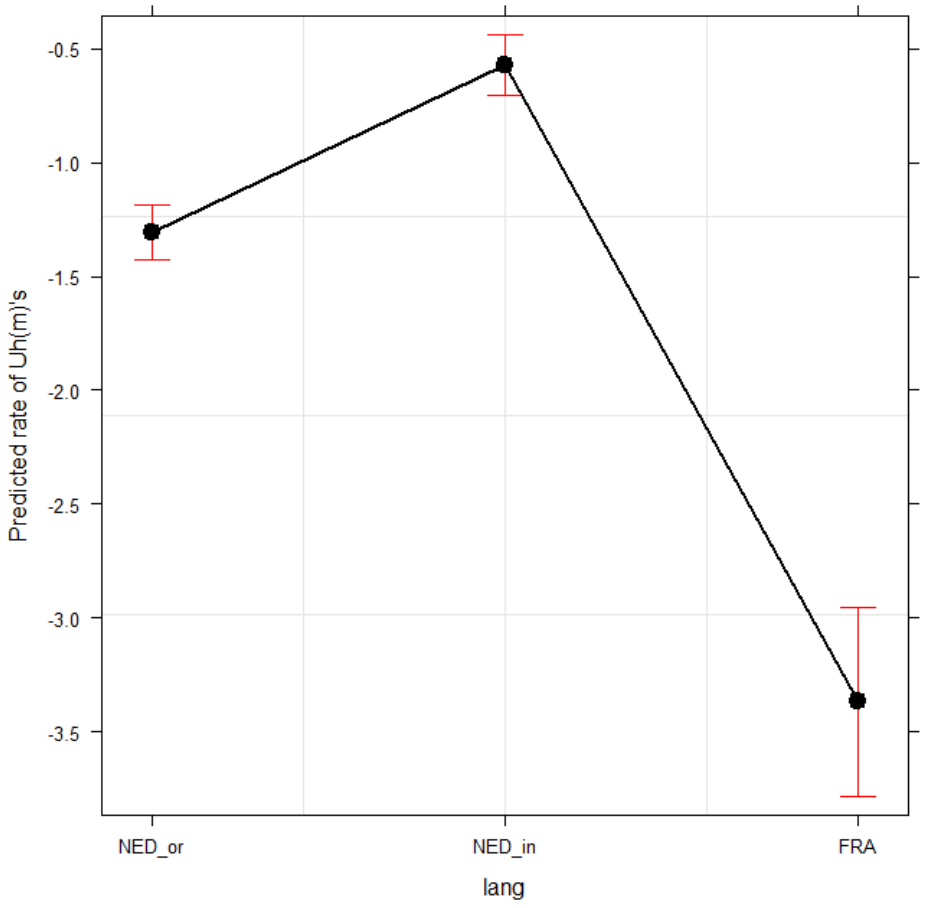
- Lexical density
- Proportion of numbers

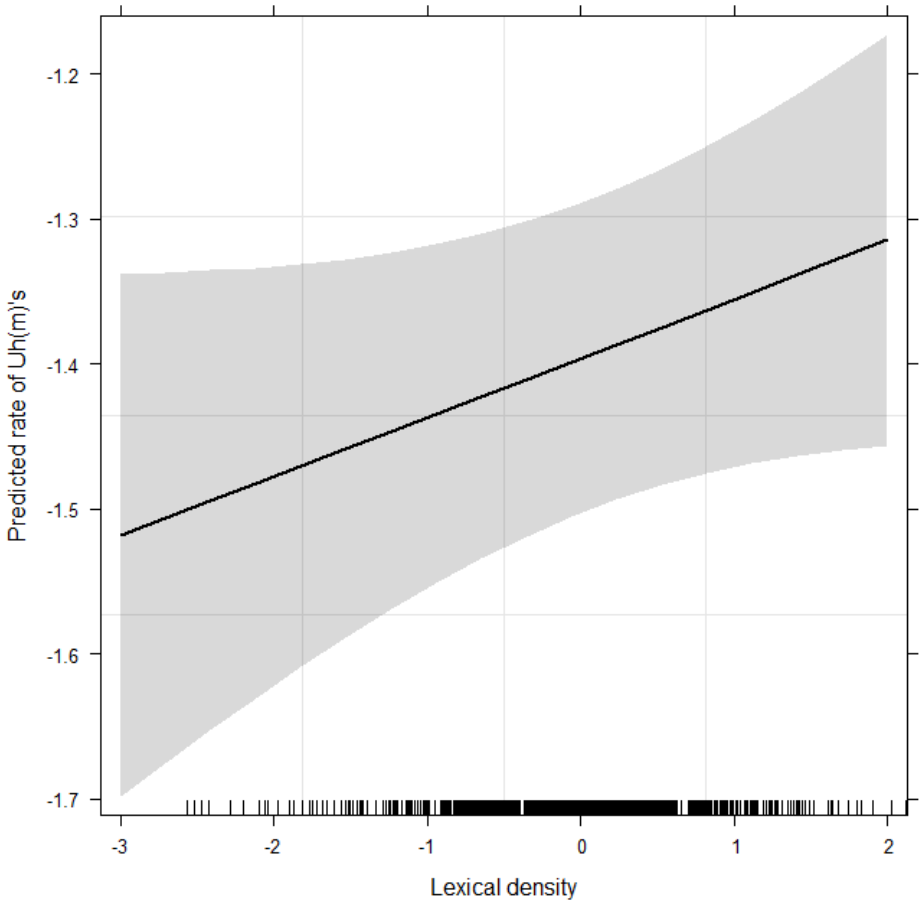
Second-level predictor:

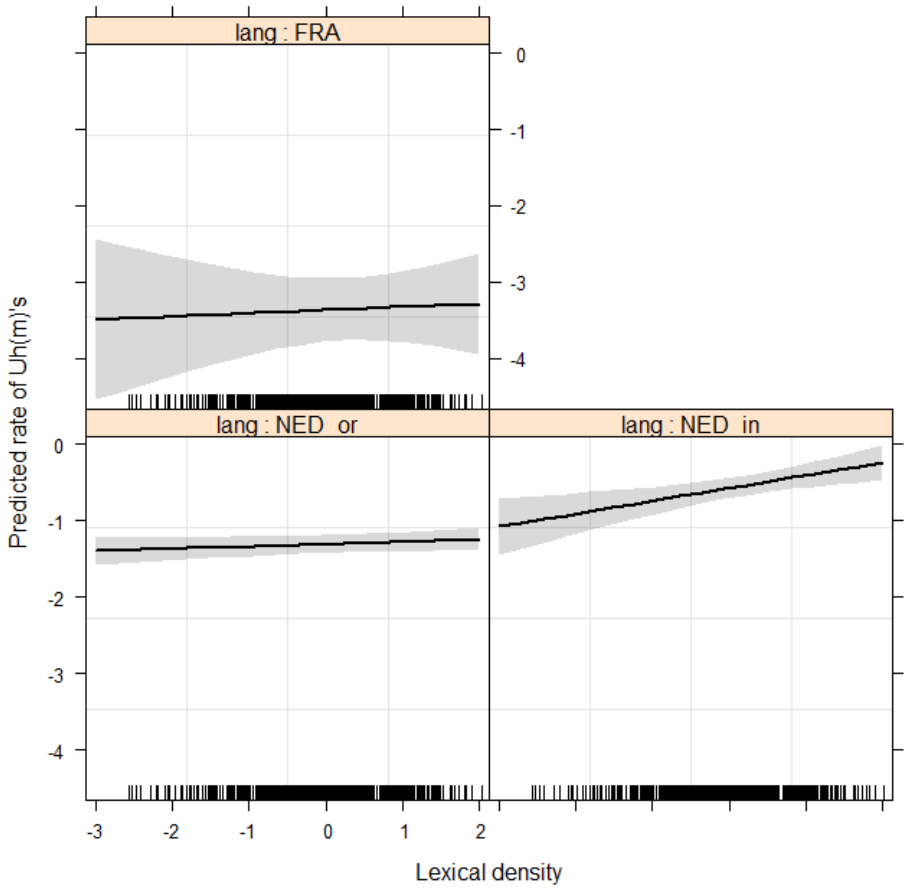
- Delivery rate

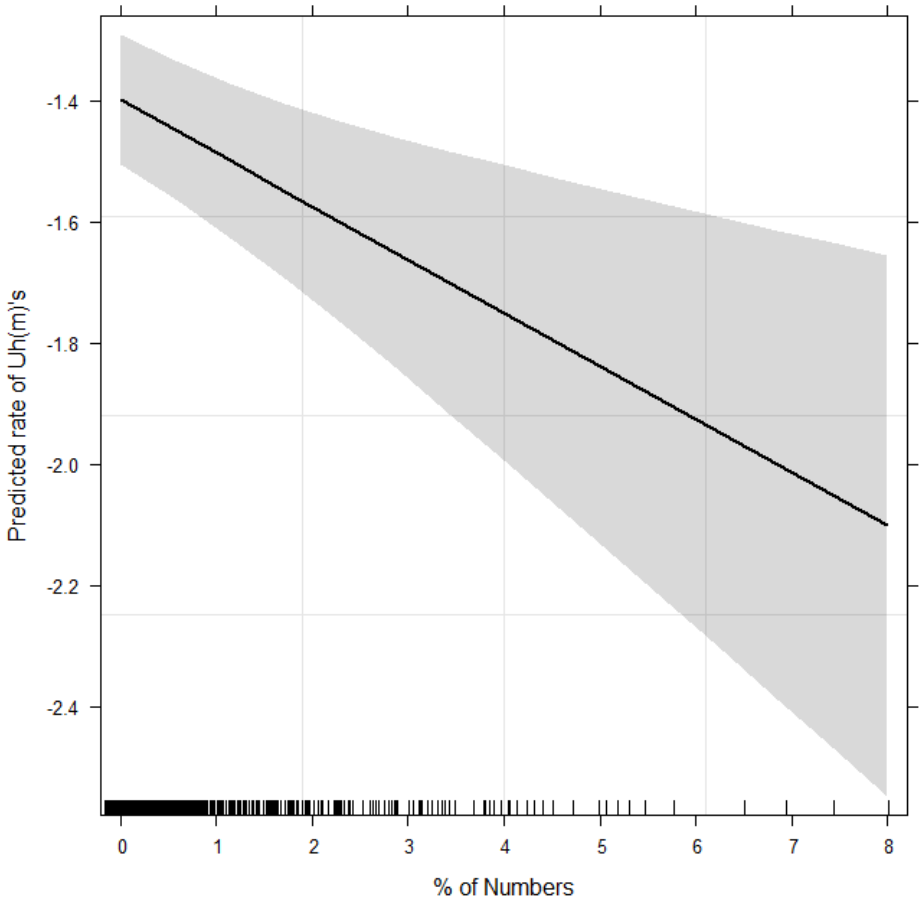
Analysis

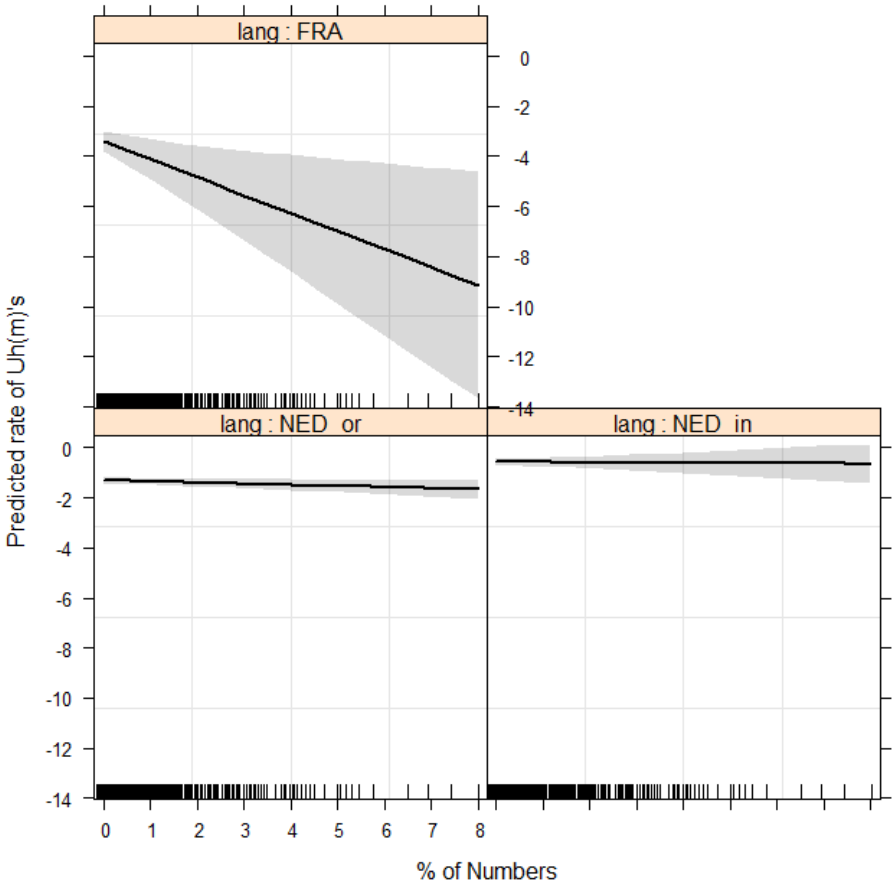
1. ‘Bakerian’ comparison of the output in the 3 ‘languages’ (*NED_or*, *NED_in* & *FRA*)
2. Comparison of the input and output in the interpretations

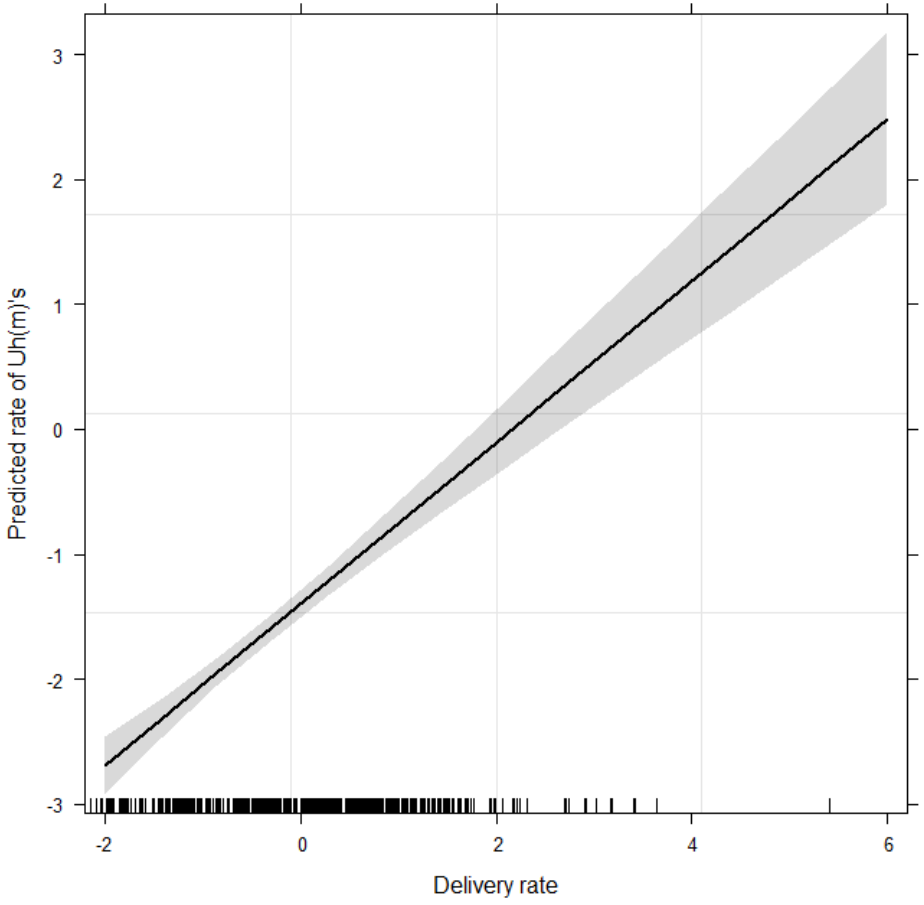

-0.5
-1.0
-1.5
-2.0
-2.5
-3.0
-3.5
Predicted rate of Uh(m)'s
NED_or
NED_in
FRA
lang

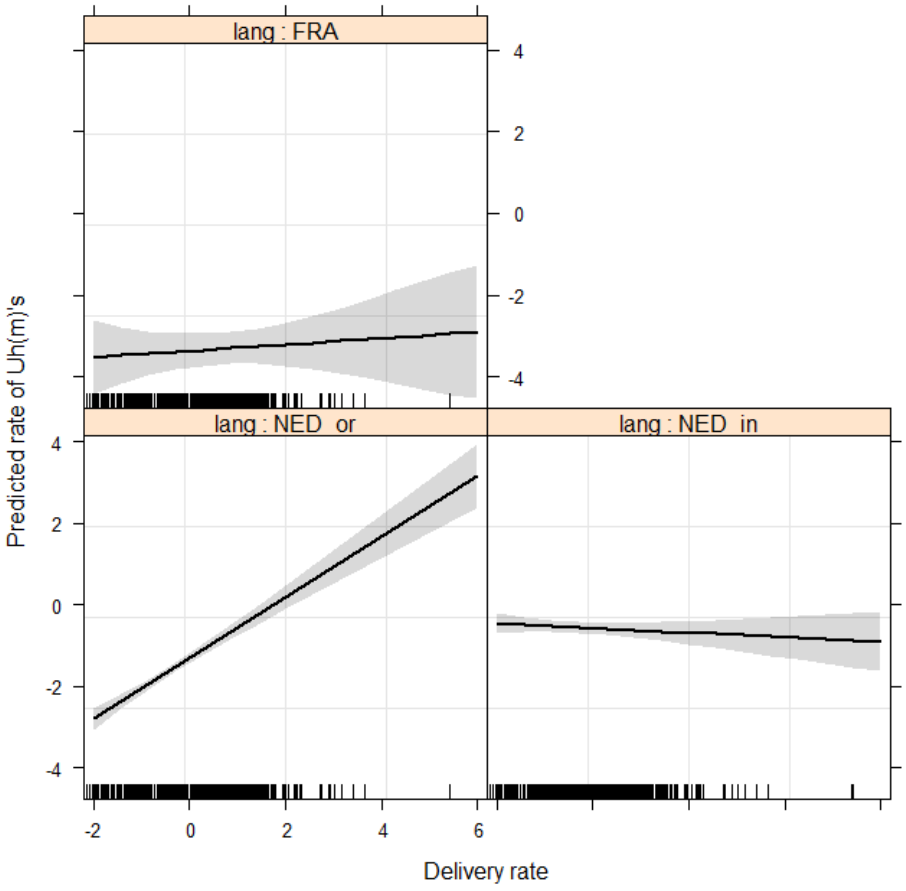

Predicted rate of Uh(m)'s
-1.2
-1.3
-1.4
-1.5
-1.6
-1.7
-3
-2
-1
0
1
2
Lexical density


lang : FRA
lang : NED or
lang : NED in
Predicted rate of Uh(m)'s
0
-1
-2
-3
-4
-3
-2
-1
0
1
2
Lexical density


Predicted rate of Uh(m)'s
-1.4
-1.6
-1.8
-2.0
-2.2
-2.4
0
1
2
3
4
5
6
7
8
% of Numbers

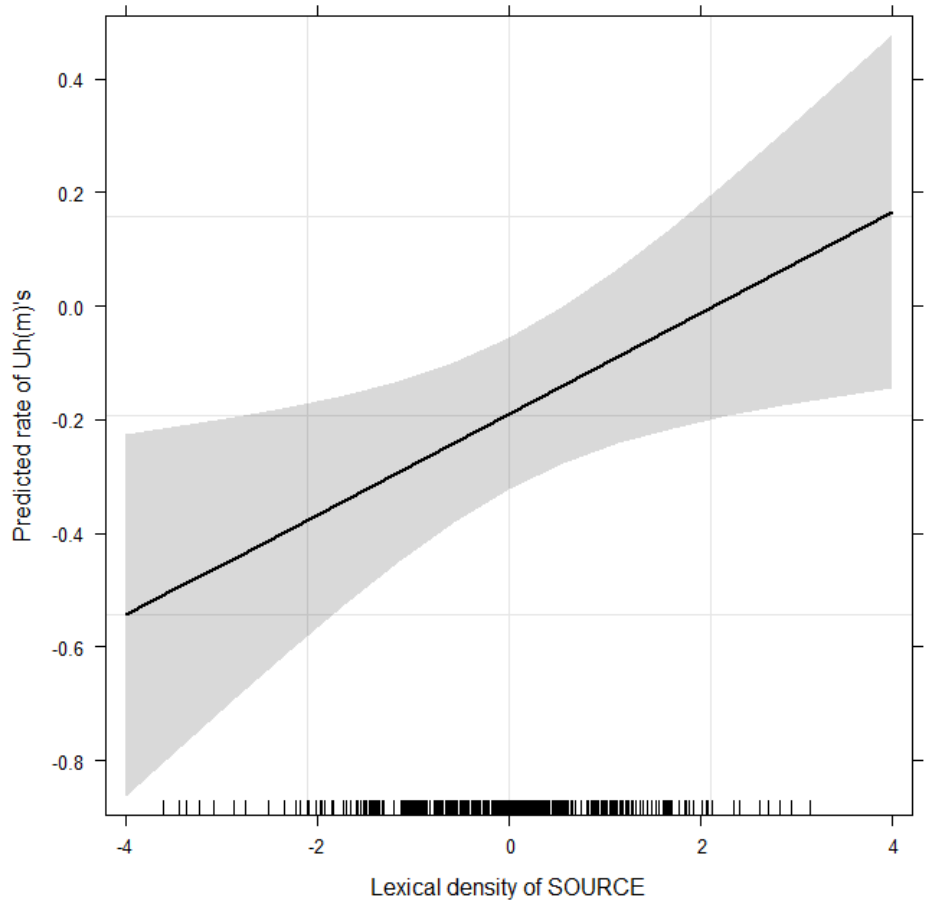

lang : FRA
lang : NED or
lang : NED in
Predicted rate of Uh(m)'s
0
-2
-4
-6
-8
-10
-12
-14
0
1
2
3
4
5
6
7
8
% of Numbers

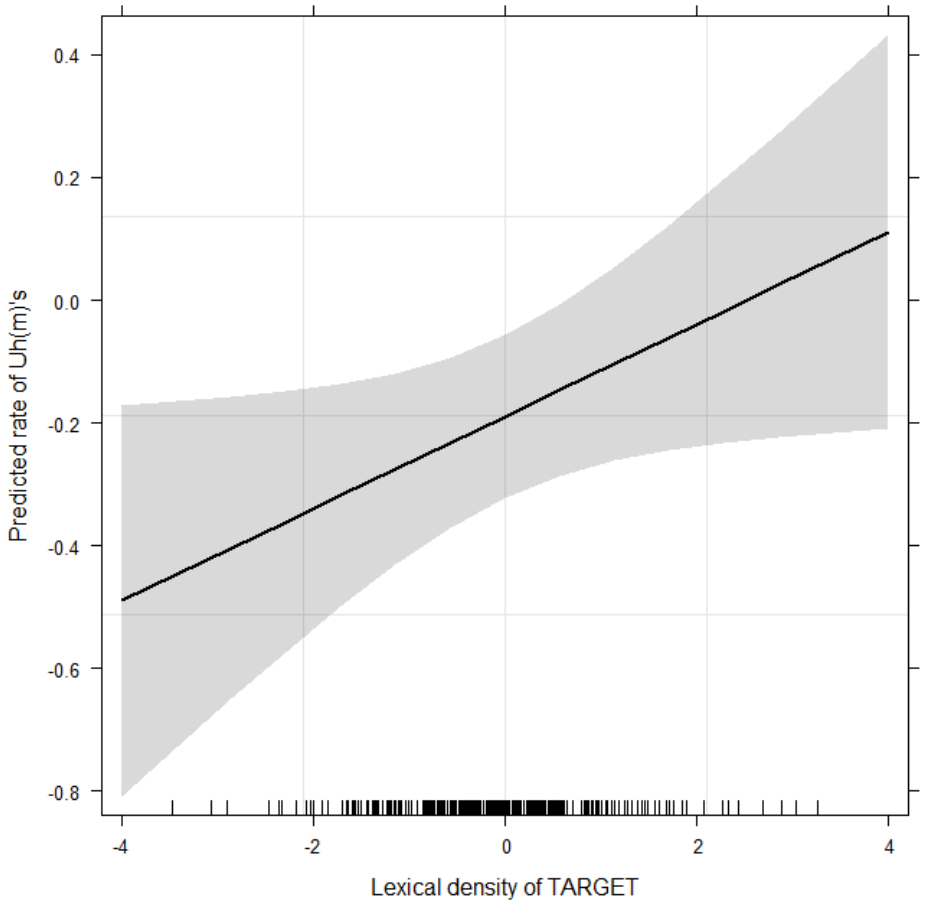

Predicted rate of Uh(m)'s
Delivery rate
3
2
1
0
-1
-2
-3
-2
0
2
4
6

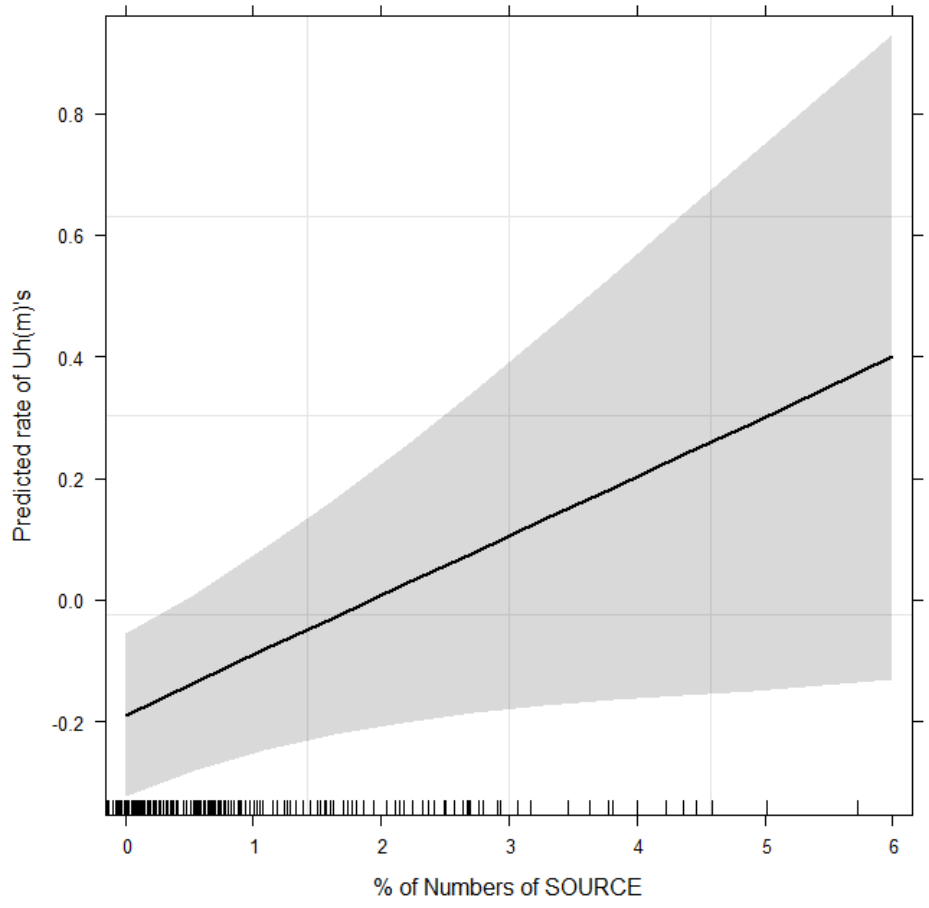

lang : FRA
lang : NED or
lang : NED in
Predicted rate of Uh(m)'s
Delivery rate
4
2
0
-2
-4
-2
0
2
4
6

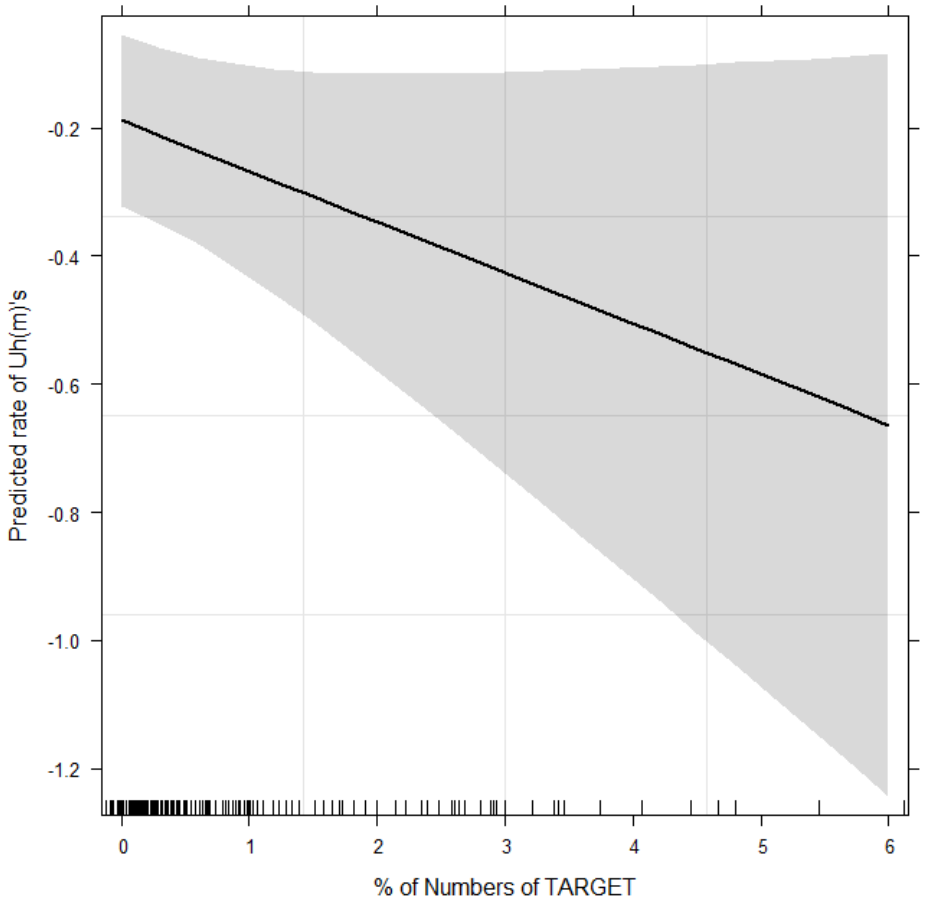
Analysis

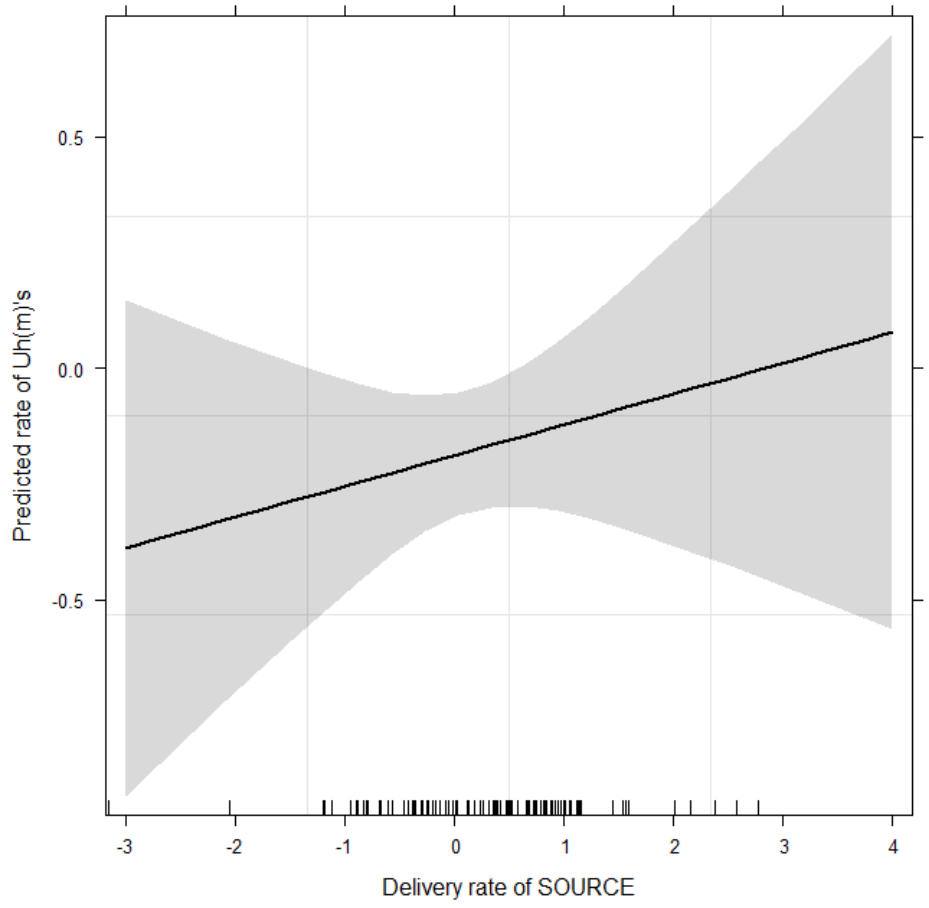
1. ‘Bakerian’ comparison of the output in the 3 ‘languages’ (*NED_or*, *NED_in* & *FRA*)
2. Comparison of the input and output in the interpretations

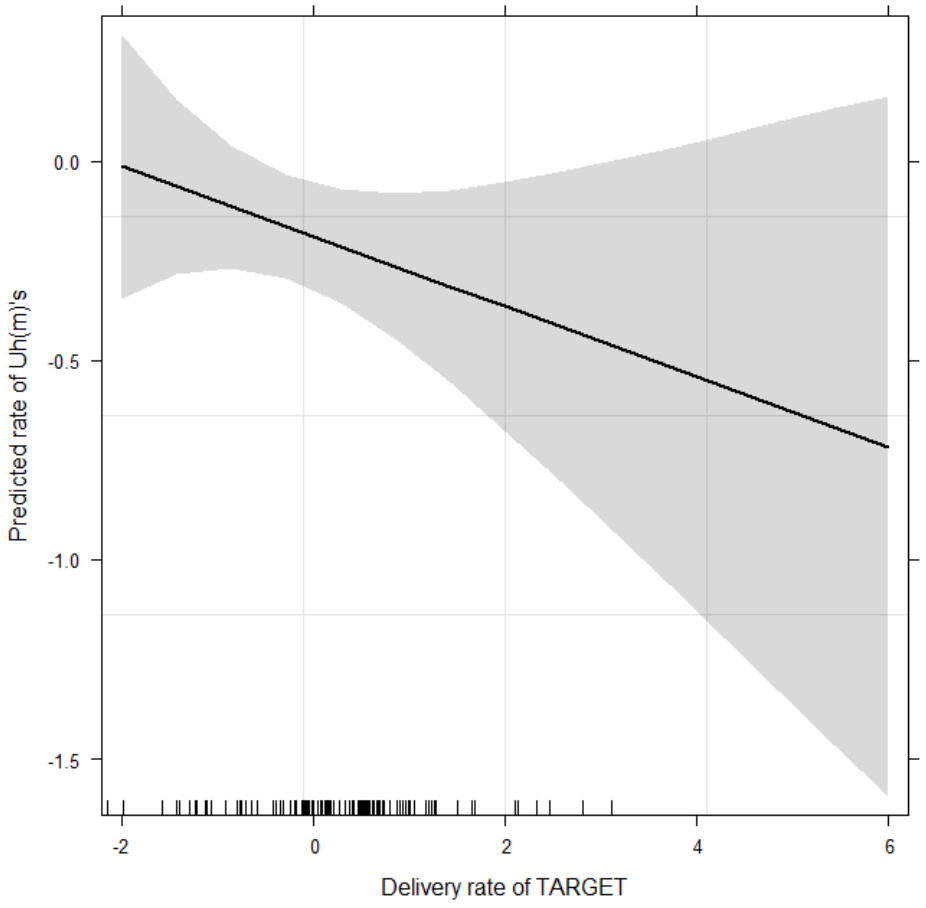

Predicted rate of Uh(m)'s
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-4
-2
0
2
4
Lexical density of SOURCE


Predicted rate of Uh(m)'s
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-4
-2
0
2
4
Lexical density of TARGET

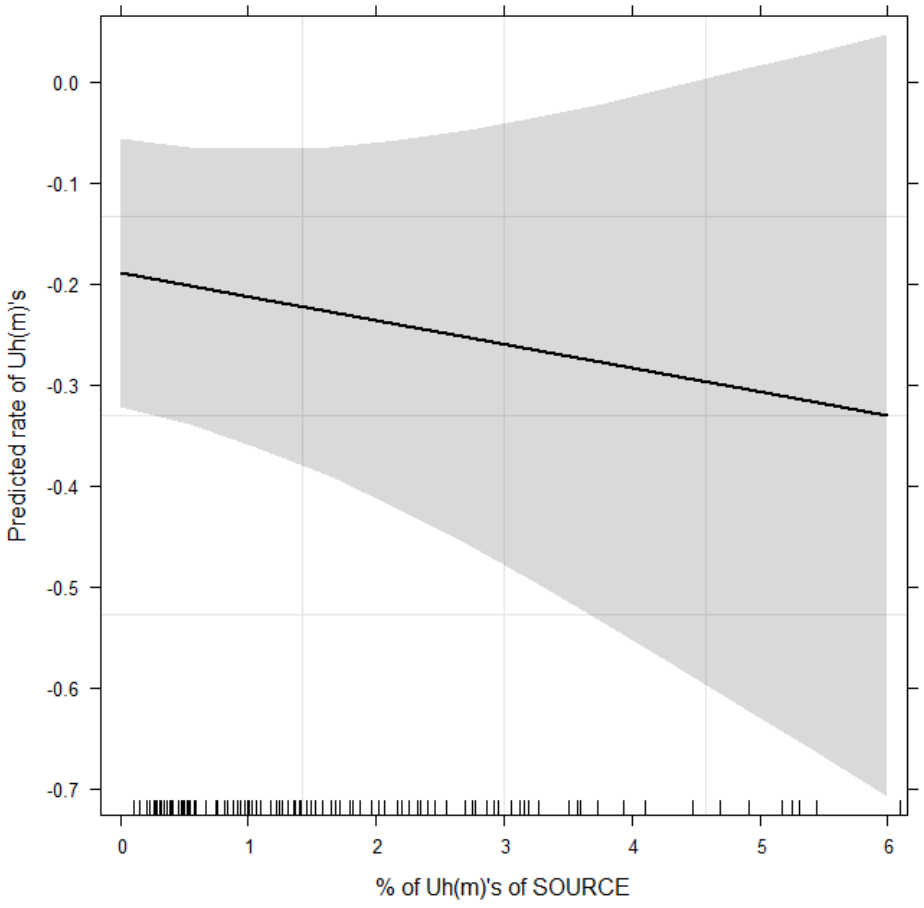

Predicted rate of Uh(m)'s
0.8
0.6
0.4
0.2
0.0
-0.2
0
1
2
3
4
5
6
% of Numbers of SOURCE


Predicted rate of Uh(m)'s
-0.2
-0.4
-0.6
-0.8
-1.0
-1.2
0
1
2
3
4
5
6
% of Numbers of TARGET


Predicted rate of Uh(m)'s
0.5
0.0
-0.5
-3
-2
-1
0
1
2
3
4
Delivery rate of SOURCE


Predicted rate of Uh(m)'s
0.0
-0.5
-1.0
-1.5
-2
0
2
4
6
Delivery rate of TARGET

Cf. Goldman-Eisler (1967); Gerver (1975)



Conclusion

Interpreters produce more *uh(m)*'s than non-interpreters

Rate of *uh(m)* in interpreting is enhanced by lexical density and numbers in the source

In non-interpreting, *uh(m)* is more related to the delivery rate

Conclusion

Negative effect of numbers in the target may point to omission

Prospect of applying Semantic Vector Spaces to interpreting

Underlines advantages of naturalistic corpus data

Thank you!

koen.plevoets@ugent.be
bart.defrancq@ugent.be